

EMBODIED SCHEMAS FOR CROSS-MODAL MAPPING IN THE DESIGN OF GESTURAL CONTROLLERS

Mark Linnane, Linda Doyle & Dermot Furlong

A conceptual framework for the design of intuitive gestural controllers for timbre manipulation from an Embodied Cognition perspective is proposed. This framework incorporates corporeal mimesis, affective/kinaesthetic dynamics, image schematic and conceptual metaphoric organization in order to speculate on possibilities in the design of mapping schemes between interface and synthesis algorithm.

Introduction

In contrast to traditional musical instruments, digital interfaces introduce an “arbitrary” component by encoding physical gesture as an electronic signal. This arbitrary component interrupts the natural, ‘felt’ sense of the interaction between performer and instrument and necessitates the introduction of control metaphors and mapping schemes that are not necessarily readily intelligible for non-technical users.

This paper focuses on the design of interfaces for the control of synthesized timbre. Several mapping strategies for control of synthesis have been proposed. However, it has been noted that empirical investigation of the “natural tendencies for multimodal mapping” is required in order to elaborate a generalizable set of cross-modal mappings. [1] We propose that the design of transparent, intuitive interfaces for timbre manipulation can be grounded in empirical analysis and subsequent artistic practice-based research focussed on cross-modal correlations between the user's embodied experience and a timbral perceptual domain. Gestural control is suggested as an appropriate paradigm in this context because it offers the possibility of affording intuitive control of sound via an intermediate ‘embodied’ mapping scheme based on ‘embodied’ characterizations of timbre and gesture.

Embodied Cognition offers a promising conceptual basis for this research. We suggest that structures common to cognition, multimodal perception and physical gesture can be identified. These correlations may then form the basis of a generalizable set of cross-modal mappings to be incorporated in the design of a gestural controller for timbral manipulation. This paper is a speculative outline of some possibilities for the application of embodied cognition in the design of gestural controllers for timbre manipulation.

Embodied Cognition

BACKGROUND

Embodied Cognition is an emerging research strand in a number of different disciplines that posits embodiment as profoundly constitutive of cognition. In general, Embodied Cognition analyses support the view of cognitive processes as creative, figurative acts structured by the teleological interactions of embodied agents with their environments and cultures rather than as largely passive processes of repre-

sentation and calculation. [2] A central claim of Embodied Cognition is that cognition and action are inextricably linked in lived experience. [3] Significantly, studies suggest that the specific structures of an organisms perceptual apparatus contribute to a structuring of all cognitive processes, not just those specifically recruited in perception.

Embodied Cognition analyses reveal a number of common primal structures and processes in cognition. These are learned via the common human experience of having a physical body and interacting in physical and cultural environments. Affective/kinaesthetic dynamics, Sensorimotor Mimesis, Image Schemata and Conceptual Blending have been identified for empirical analysis in this study. Essentially, the research question here is, for example, are Image Schemata activated by judgments of timbral quality and by gesture production and recognition? Can such schema be reliably demonstrated empirically? If so, what correlations exist, if any, between schema activated by timbre and those activated by gesture?

IMAGE SCHEMATA

Image schemata are basic features in cognitive processing acquired in early childhood development. They are based on basic dynamic embodied patterns and they give structure to more abstract conceptual processes. They are non-propositional, pre-conceptual structures that have a role in organising mental representations into coherent, meaningful units. They have a basic logic and are implicated in the formation of new concepts. Significantly for this study, they are focussed on perceptual interactions, and bodily and linguistic experience in social and historical contexts and they are inherently meaningful because of this embodied grounding. They are neurally encoded. [4]

The following are three simple examples of Image Schemata. The Link Schema is a simple physical or metaphorical structure whereby one object is linked to, and constrained by, another according to a basic symmetrical logic. Examples are the concept of a causal “connection”, or the physical act of connecting a laptop to its power source. The Container Schema has an Interior-Boundary-Exterior structure. It is based on the bodily experience of being a container and of being contained by something. The basic logic is that of the mutual exclusivity of interior and exterior and the necessity of being in one state or the other. The Source-Path-Goal Schema is based on bodily experiences such as throwing a ball. Other examples of Image Schemata as suggested by Johnson [5] are part-whole, centre-periphery, link, cycle, interaction, contact, pushing, balance, near-far, front-back and above-below.

An analysis of concepts of physical relation in the sentence, ‘The book is on the desk’ [3] illustrates the general form in which Image Schemata are implicated in the construction of meaning in linguistic expressions. Meaning here emerges by virtue of the fact that the Image Schemata activated by the linguistic expression are derived from a composite of neurally encoded embodied experiences such as being above, being in contact with, supporting, and so on. These experiences activate specific schemata which give meaning to the word ‘on’ in this expression. The experience of being ‘above’ yields an orientational schema; ‘contact’ yields a topological schema; and ‘support’, a force-dynamic schema. An important aspect of this process is that meaning does not arise through some kind of correlation between concept and objective ‘fact’, or by correlation of symbol and referent, but through a composite of activated schemata.

It is significant that several of these schemata have a structural identity with gestures or gestural components (e.g. source-path-goal schema, centre-periphery, above-below, pushing etc.). Johnson also presents analysis of image schemata across visual, kinesthetic and auditory modalities. [5] Image Schemata,

therefore, may offer a means of organizing gestural analysis in terms of the embodied structures of perception in other modalities.

CONCEPTUAL BLENDING

Lakoff and Johnson [6] argue that there are no literal definitions, only trans-domain mappings via metaphor. New concepts or categories are formed via the conceptual blending or mapping of a set of schemata from one domain to a target domain. For example, conceptual blending based on the Source-Path-Goal Schema yields complex causal patterns, such as that of ice changing state as it melts. Metaphorical conceptualization is, therefore, significantly constitutive of all thought. It is systematic, fundamental to language and thought and embodied. Such conceptual blends are implicated in cross-modal correlations such as judgments of timbre as being, for example, rounded, harsh, warm, heavy etc. It is important to note that gesture is fully characterizable within the framework of conceptual metaphoric mappings across modalities. [7]

SENSORIMOTOR MIMESIS

Sensorimotor Mimesis is any of several means by which humans imitate, consciously or unconsciously, covertly or overtly, an environmental or social stimulus. It is a crucial process in the psychological development of infants. It is often co-activated with other behaviors such as speaking, listening to music and working out logical problems. Mimesis is particularly significant because it appears to be implicated in cognitive tasks such as understanding and in affective responses to stimuli. In particular, vocal mimesis seems to be an important part of the music listening experience. People have the tendency to move in sympathy with music and, it appears, that this behavior is fundamentally constitutive of our understanding of, and affective response to music.

AFFECTIVE/KINAESTHETIC DYNAMICS

There is a formal congruency between motion and emotion. The felt quality of emotion is grounded in demonstrable dynamic patterns of physical expression. This is a reciprocal dynamic, emotions are shaped by motor attitudes, just as physical movement expresses emotion. Empirical studies by Bull [8] demonstrate the inability of subjects to experience particular emotions while adopting postural attitudes considered to be antithetical. Emotional expression is a kinetic phenomenon that has spatiality, temporality, intensity, and manner of execution. Phenomenological bracketing has been used to elucidate the dynamic structure underlying these forms. *Tensional* quality is mediated by the felt effort of postural attitude. *Linear* quality is directional contour of movement. *Amplitudinal* quality is the amount of extensiveness or constrictiveness of a posture. *Projectional* quality is the manner in which energy is released.

Sensorimotor Mimesis, considered along with Affective/Kinaesthetic Dynamics give us a theoretical framework within which analyse corporeal responses to stimuli. Mimetic responses to music are, seemingly, of a piece with our affective, aesthetic response. Affective/Kinaesthetic dynamics offer a structure whereby affective response can be assessed.

Hypothesis: Embodied Schemas for Gestural Manipulation of Timbre

In light of the above, we suggest that responses to timbre stimuli may be characterized as mediated by corporeal mimetic engagement according to affective/kinaesthetic dynamic structures. Timbres 'feel' a certain way due to this corporeal engagement. We further suggest that judgments of difference or similarity between timbres is mediated by an Image Schematic/Metaphoric structuring of differing corporeal attitudes. We suggest that such mapping processes mediate meaningful 'navigation' of timbre space via physical gesture. This view is supported by the observation that natural linguistic descriptions of timbre tend to emphasize embodied, cross-modal mappings through the use of multi-modal, embodied descriptors.

We suggest that empirical analysis may show that identical Image Schemata are activated by motor, visual and auditory stimulation. Substantial trends in cross-modal association have already been demonstrated. [9] We speculate that common cognitive structures may be shown to account for these correlations and to account for the tendency to describe sound in terms of weight, force, speed, intensity, emotion, spatial position and orientation, containedness, gravity, density, amplitude, colour, order, chaos etc. Essentially, we maintain that identical Image Schemata structure perception across modalities, thus allowing for perceptions of close correlation between certain gestures or shapes for example and certain timbres.

Such correlations would present a grounded basis for mapping schemes in gestural interfaces for sound synthesis. We propose that an empirical study of gesture performance and timbre perception that attempts to find the basis for cross-modal mappings between each is the first step in the design of these mappings for synthesis interfaces.

Design Approach

We propose empirical analysis and subsequent artistic practice-based research in order to elucidate cross-modal patterns. Study 1 will take the form of a preliminary subjective study designed to assess Image Schema activation in listeners in response to audio presentations of synthesized timbres and also, to assess the methodology with a view to later tests. The intent here is to assess the viability of a methodology which assumes that the perception of timbre difference is structured according to Image Schematic principles. Some of the following schema will form the test set; Path, Source-Path-Goal, Center-Periphery, Compulsion, Attraction, Link, Scale, Equilibrium, Full-Empty, Near-Far, Mass-Count, Iteration, Above/Below, Vertical Orientation, Length (extended trajectory), Rough-smooth/Bumpy-smooth.

Subjects will be presented with a set of synthesized timbres. These will be designed to include examples of pitched tones, percussive tones, instrument models and novel timbres. Each timbre example will be presented in pairs, once in an unmodified fashion then with one modification. Timbres will be modified in terms of some of the well-established factors known to be involved in timbre e.g. spectral envelope, harmonicity, spectral centroid, attack and decay envelopes.

A number of methods for assessing Image Schema activation are currently being considered. Subjects may be asked to supply linguistic descriptions of the presentations. The descriptions will be constrained to a limited set of options or to a defined metric and chosen in an onscreen interface. Subjects may also be asked to select the most appropriate graphical representation of an Image Schematic structure that describes the relationship between pairs of presented timbres. Synthesis will be implemented in

Max/MSP as will the user interface and response data collection system. Data will be subjected to statistical analysis in order to identify trends, if any, in subject response.

Study 2 will incorporate gesture capture via inertial sensors and a pair of sensor-enabled gloves. These sensors will enable the tracking of the relative position of the arms and the amount of contraction/expansion in the arms. Subjects will be asked to adopt whatever gesture or gestures they feel are appropriate responses to timbre stimuli as presented in the pilot study. Subjects will also be asked to perform gestures in response to linguistic and pictorial presentations of image schema as detailed above. This test will yield some measure of the corporeal mimetic response to timbre presentations in that it will provide some data on the Linear, Amplitudinal and Projectional qualities of the affective/kinaesthetic response. It will also establish whether it is plausible to assess schema activation in gesture production.

Study 3 will be a test of prototype mappings. If Tests 1 and 2 yield clear data, then there may be correlations between kinaesthetic and schematic activations in the case of both timbre and gesture that can form the basis of the design of systems in a phase of artistic practice-based tests focussing on art installation and music performance contexts.

Conclusion

Embodied Cognition is an area of cognitive science that focusses on perception, cognition and action as profoundly shaped by the human experience of having a body and living in a physical environment and in a culture. It suggests an alternative approach that complements current initiatives in the design of interactive technologies and gestural controllers. Embodied Cognition offers a novel way to analyse the complex interactions between user and technology in terms of the fundamental categories of embodied existence. These categories are involved in organizing perceptual phenomena into coherent, meaningful units and are implicated in the formation of new concepts. In turn, the design of meaningful controllers depends upon empirical knowledge of the fundamental categories whereby human beings interact with and understand their environment. In using these structures as a conceptual basis, it may be possible to identify correlations between gesture and phenomena, such as timbre, in other modalities. Such correlations are the necessary building blocks of a grounded cross-modal mapping schema on which to base the design of controllers that allow for meaningful gestural interaction with sound.

References and Notes:

1. P. Maes, M. Leman, M. Lesaffre, M. Demey and D. Moelants, "From expressive gesture to sound: The development of an embodied mapping trajectory inside a musical interface," *Journal on Multimodal User Interfaces* 3, nos. 1–2 (2010): 67–78.
2. M. Imaz and D. Benyon, *Designing with blends* (Cambridge, MA: The MIT Press, 2007).
3. G. Lakoff and R. Núñez, *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being* (New York: Basic Books, 2000).
4. T. Rohrer, "Image Schemata in the Brain," in *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, ed. B. Hampe, 165–196 (Berlin: Mouton de Gruyter, 2006).
5. M. Johnson, *The Body in the Mind: the bodily basis of meaning, imagination, and reason* (Chicago, IL: Chicago University Press, 1987)
6. G. Lakoff and M. Johnson, *Philosophy In The Flesh: the Embodied Mind and its Challenge to Western Thought* (New York: Basic Books, 1999).
7. R. Núñez, "A Fresh Look at the Foundations of Mathematics: Gesture and the Psychological Reality of Conceptual Metaphor," in *Gesture and Metaphor*, eds. A. Cienki & C. Müller, 93–114 (Amsterdam: John Benjamins, 2008).
8. N. Bull, *The attitude theory of emotion* (New York: Nervous and mental disease monographs, no. 81, 1951).
9. J. Simner, C. Cuskley and S. Kirby, "What sound does that taste? Cross-modal mappings across gustation and audition," *Perception* 39 (2010): 553–569.